5th Asia Pacific Conference on Industrial Engineering and Operations Management
Tokyo, Japan, September 10-12, 2024
Publisher: IEOM Society International, USA DOI: 10.46254/AP05.20240088
Published: September 10, 2024

Enhancing Robustness in Explainable AI: Applications in Community Renewable Energy Engagement

Firouzeh Rosa Taghikhah
Business School
University of Sydney
Camperdown, 2050, NSW, Australia
Firouzeh.Taghikhah@sydney.edu.au

Philip Cai
School of Computing
Australian National University
Canberra, 0200, ACT, Australia
Anzhe.Cai@anu.edu.au

Ivan Bakhshayeshi
School of Biomedical Engineering,
University of New South Wales
Sydney, 2032, NSW, Australia.
I.Bakhshayeshi@uns.edu.au

Masoud Taghikhah
Faculty of Computer Science
Dresden University of Technology
Dresden, 01069, SN, Germany
m.taghikhah@gmail.com

Abstract

The robustness of explainable artificial intelligence (XAI) models is a cornerstone for their effectiveness and trustworthiness in high-stakes environments. As AI technologies permeate sensitive areas such as environmental monitoring and energy management, the demand for models that not only perform consistently but also offer transparent decision-making processes intensifies. This study introduces a transformative methodological innovation in XAI that significantly enhances the robustness and interpretability of predictive models. Employing advanced machine learning algorithms—XGBoost, Multi-layer Perceptron, and Keras Sequential Neural Networks—integrated with leading-edge XAI tools like Tree SHAP, Kernel SHAP, and Deep SHAP, our framework sets a new standard in the fidelity and accountability of AI systems.Central to our methodology is the development of two novel metrics: the Stability Correlation Index (SCI) and the Explanation Integrity Metric (EIM). These metrics are specifically designed to critically evaluate the stability and reliability of explanations provided by AI models across varied and unpredictable data landscapes. By quantifying the consistency of model explanations under dynamic conditions, these metrics address a crucial gap in current AI practices—ensuring that AI systems remain robust against data perturbations and environmental changes, which is essential for their deployment in real-world scenarios.

Our application of this robust framework to a case study in Community Renewable Energy (CRE) illustrates its practical value. In CRE projects, understanding the multifaceted factors influencing community engagement is vital for the successful implementation and sustainability of renewable initiatives. By applying our robust XAI methods, we not only enhance the predictive accuracy but also provide stakeholders with clear, interpretable insights into the complex dynamics of community participation. This dual capability is indispensable for designing effective, data-driven strategies that promote widespread adoption and support for renewable energy projects.

Keywords
Explainable Artificial Intelligence, Machine Learning Robustness, Stability Metrics, Community Renewable Energy, Predictive Analytics, Transparency in AI

Biographies
Dr. Firouzeh Rosa Taghikhah is a lecturer at the Business School, University of Sydney, specializing in business analytics with a keen interest in decision analysis and causal discovery. Her research focuses on developing advanced methodologies for explainable AI, striving to enhance transparency and accountability in decision-making processes within various business contexts.

Philip Cai is currently pursuing his Master's degree at the School of Computing, Australian National University. His research interests center around the field of explainable artificial intelligence (XAI), where he explores methodologies to improve the interpretability and reliability of machine learning models, particularly in contexts requiring robust decision support systems.

Ivan Bakhshayeshi is a researcher at the School of Biomedical Engineering, UNSW Sydney. His work focuses on the development of explainable artificial intelligence (XAI) algorithm within the healthcare sector. Ivan is dedicated to advancing XAI technologies to improve diagnostic accuracy, treatment personalization, and patient outcomes. His research aims to bridge the gap between complex AI algorithms and clinical practice, ensuring that healthcare professionals can understand and trust AI-driven decisions, thus enhancing both patient care and medical research.

Masoud Taghikhah, a recent graduate from the Faculty of Computer Science at Dresden University of Technology, has a profound interest in the applications of explainable AI and outlier detection. His work aims to bridge the gap between technical AI implementations and practical applications, ensuring that AI technologies are both accessible and comprehensible to users across diverse industries.